

PUERTO RICO AND VIRGIN ISLANDS  
PRECIPITATION FREQUENCY PROJECT

Update of *Technical Paper No. 42* and *Technical Paper No. 53*

Nineteenth Progress Report  
1 January 2005 through 31 March 2005

Hydrometeorological Design Studies Center  
Hydrology Laboratory

Office of Hydrologic Development  
U.S. National Weather Service  
National Oceanic and Atmospheric Administration  
Silver Spring, Maryland

April 2005

### DISCLAIMER

The data and information presented in this report are provided only to demonstrate current progress on the various technical tasks associated with this project. Values presented herein are NOT intended for any other use beyond the scope of this progress report. Anyone using any data or information presented in this report for any purpose other than for what it was intended does so at their own risk

## TABLE OF CONTENTS

1. Introduction .....	1
2. Highlights .....	3
3. Progress in this Reporting Period.....	5
4. Issues.....	23
5. Projected Schedule and Remaining Tasks.....	24
References.....	25

## PUERTO RICO AND VIRGIN ISLANDS PRECIPITATION FREQUENCY PROJECT

Update of *Technical Paper No. 42* and *Technical Paper No. 53*

### 1. Introduction

The Hydrometeorological Design Studies Center (HDSC), Hydrology Laboratory, Office of Hydrologic Development of NOAA's National Weather Service is updating its precipitation frequency estimates for Puerto Rico and the Virgin Islands. Current precipitation frequency estimates for the area are contained in *Technical Paper No. 42* "Generalized estimates of probable maximum precipitation and rainfall-frequency data for Puerto Rico and Virgin Islands" (U.S. Weather Bureau, 1961) and *Technical Paper No. 53* "Two- to ten-day rainfall for return periods of 2 to 100 years in Puerto Rico and Virgin Islands" (Miller, 1965). The new project includes collecting data and performing quality control, compiling and formatting datasets for analyses, selecting applicable frequency distributions and fitting techniques, analyzing data, mapping and preparing reports and other documentation.

The project will determine annual precipitation frequencies for durations from 5 minutes to 60 days, for average recurrence intervals from 1 to 1,000 years. The project will review and process all available rainfall data for the Puerto Rico and Virgin Island project area and use accepted statistical methods. The project results will be published as a Volume 3 of NOAA Atlas 14 on the internet (<http://www.nws.noaa.gov/ohd/hdsc>) with the ability to download digital files.

The project area covers Puerto Rico and the U.S. Virgin Islands of St. Thomas, St. John and St. Croix. The project area is currently divided into 6 regions for analysis (Figure 1).

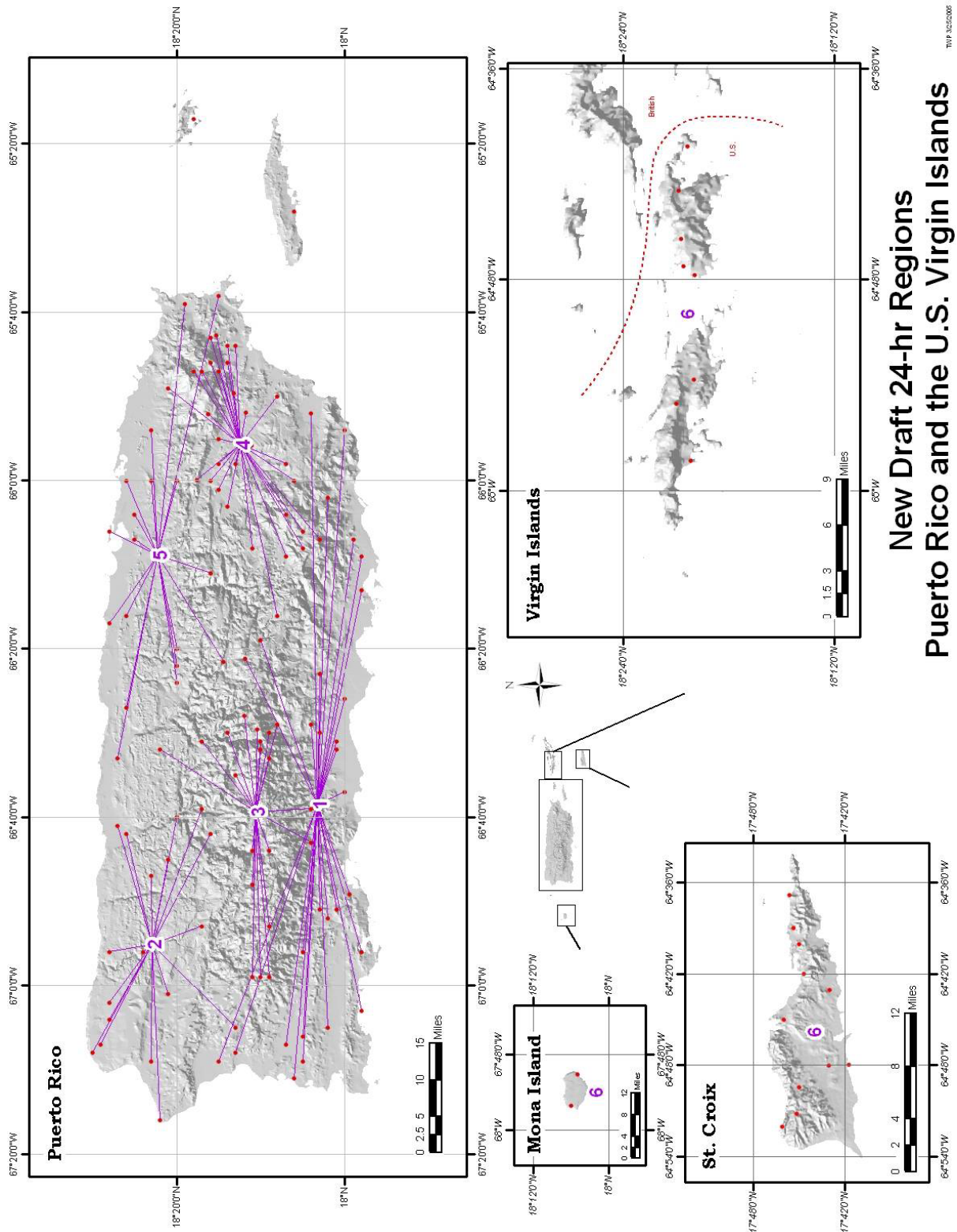


Figure 1. Puerto Rico Precipitation Frequency project area and the 6 regional groups.

## 2. Highlights

15-minute, 1-hour and 1-day data were quality controlled using the new spatially-based quality control procedure, *QCseries* that screens annual maximum series (AMS) and partial duration series (PDS) to identify maximum precipitation values that are suspect relative to concurrent data at nearby stations. Additional information is provided in Section 3.1.1, Data Series Quality Control; Data Collection and Quality Control.

After a thorough investigation, it was decided not to use the 15-minute dataset from the United States Geological Survey (USGS) due to a short period of record and some data quality issues. Automated Surface Observing System (ASOS) data through 2003 were added to the n-minute dataset and are being quality controlled. Hourly data through 2004 were not available to add to the dataset. Initial hourly data quality control is complete. Additional information is provided in Section 3.1.2, 15-minute, N-minute, and Hourly Data Collection and Quality Control.

Daily data through 2004 were quality controlled and added to the daily dataset. A check of all low 1-day annual maxima was completed resulting in eleven data corrections. Accumulated daily data during hurricane-type events were investigated and objectively distributed over their days of accumulation. As a result of this more appropriate distribution of hurricane-type events, mean annual maxima at stations increased an average of 3.3%. Cases where 1-day monthly maxima observations were significantly lower than concurrent co-located 24-hour accumulations were identified and resolved. Stations with large gaps (i.e., sequential missing years) in the 1-day annual maximum series of stations were flagged and examined on a case-by-case basis to ensure that the data on either side of the gap were from the same population. Several issues related to station location or data were identified and addressed. Additional information is provided in Section 3.1.3, Daily Data Collection and Quality Control.

It was decided to use 20 data years for daily data and 15 data years for hourly data as a minimum requirement for a station in the analysis. To accurately extract statistically meaningful values, it is ensured that each year has a sufficient number of data, particularly in an assigned "wet season". Preliminary wet seasons were defined using histograms of the percentage of annual maxima that occurred in each month for each region. Additional information is provided in Section 3.2, Annual Maximum Series Extraction.

Conversion factors for 1-day to 24-hour and 2-day to 48-hour were calculated. The factors are 1.208 and 1.134, respectively. Additional information is provided in Section 3.3, Conversion Factors.

Preliminary L-moments were run on the 24-hour and 1-hour data. Station coefficient of L-variation and L-skewness maps are included here. New regions were developed for the project area based on recent studies and current climatic data. The regional heterogeneity measure, H1, was calculated for the 24-hour data and modifications to

the regions are underway. Additional information is provided in Section 3.4, L-moments and Regionalization.

Preliminary temporal distributions of heavy 24-hour precipitation were calculated and graphed for the project area as a whole. Additional information is provided in Section 3.5, Temporal Distributions.

Select maps from Technical Paper 42 (U.S. Weather Bureau, 1961) were digitized to allow objective comparisons with this project's precipitation frequency estimates. Additional information is provided in Section 3.6, Spatial Interpolation.

The Precipitation Frequency Data Server (PFDS), the on-line portal for all NOAA Atlas 14 deliverables and information, underwent several important changes. The most noticeable change was the re-organization of the GIS Download page. Additional information is provided in Section 3.7, PFDS.

Progress on the development of areal reduction factors remains slow due to difficulties in completing the analysis software. Development and testing of software from the procedure described in NOAA Technical Report NWS 24 is 95% completed. Modifications have also been made in the ARF sites with respect to which stations are being used. Two statistical procedures have been prepared to test the differences between the ARF curves generated from the various sites. Additional information is provided in Section 3.8, Areal Reduction Factors.

### 3. Progress in this Reporting Period

#### 3.1 Data Collection and Quality Control

##### 3.1.1 Data Series Quality Control

Data were quality controlled using the new spatially-based quality control procedure, *QCseries* that screens data to identify precipitation values that are suspect relative to concurrent data at nearby stations. This software will improve productivity by narrowing the number of values screened manually. The *QCseries* software is still undergoing development, testing and implementation. Most importantly, an objective process is being developed to establish the criteria for computing the quality control flag for each of annual maximum and/or partial duration maximum. These flags will provide a measure of how confident the software is in the validity of the observed value. The flags are based on concurrent precipitation at nearby stations and the deviation from spatially distributed values of percent of mean annual precipitation. Although a fully optimized and proven flagging algorithm has not yet been completed, a significant portion of the logic has been determined and the software is currently functional as a quality control tool. Look for more details in future Quarterly Progress Reports.

The *QCseries* software output was evaluated for the 15-minute, 1-hour and 1-day annual maxima and partial duration time series. Quality control based on the *QCseries* output identified erroneous values linked to such factors as gage malfunctions, transmission errors, and coding errors. A total of twenty-five corrections were made to the dataset to correct erroneous values – 6 to the USGS 15-minute data, 6 to the NCDC 15-minute data, 7 to the 1-hour data, and 6 to the 1-day data.

##### 3.1.2 15-minute, N-minute, and Hourly Data

**15-minute Data.** Quality control of 15-minute observations above a threshold of 0.50 inches was completed for the United States Geological Survey (USGS) dataset. In addition, quality control using the *QCseries* software was conducted on accumulated 1-hour USGS data (for additional details see Section 3.1.1, Data Series Quality Control). The quality-controlled USGS data accumulated to 1-hour were compared with concurrent nearby National Climatic Data Center (NCDC) 1-hour stations. The means of the annual maxima of the USGS stations were consistently lower. Statistical t-tests at the 95% confidence level showed that the USGS data were not compatible with concurrent NCDC data. Therefore, because these stations had a shorter period of record and were not readily compatible with NCDC data, it was decided not to use the USGS dataset in the analyses.

**N-minute Data.** N-minute data are precipitation data measured at a temporal resolution of 5-minutes that can be summed to various “n-minute” durations (10-minute, 15-minute, 30-minute, and 60-minute). The single n-minute station in the project area, San Juan (66-8812), has a data record from January 1973 through May 1997 from NCDC.

Automated Surface Observing System (ASOS) n-minute data for the station from June 1998 through December 2004 were obtained, formatted, and added to the record. Extreme 5-minute values above a threshold of 0.5" and accumulated 1-hour values above a threshold of 1" are being checked. In addition, HDSC is attempting to find data to fill several gaps in the record.

This station will be used to generate n-minute ratios (n-min/60-min) for the 5-minute, 10-minute, 15-minute, and 30-minute durations for calculating n-minute precipitation frequency estimates. It will also be used to develop conversion factors from 1-hour to 60-minute and 2-hour to 120-minute. Given the importance of these ratios and conversion factors, time will be spent researching additional n-minute data in similar climates to supplement the n-minute data and/or validate conclusions.

**Hourly Data.** Hourly data through 2004 were not available to add at this time. The existing 1-hour dataset was further quality controlled using the *QCseries* software and a check of low annual maxima (see Section 3.1.1, Data Series Quality Control and 3.1.3, Daily Data; Low Annual Maxima for more details). No low annual maximum data problems were discovered in the 1-hour series. Also, 73 extreme hourly values  $\geq 1$ " were flagged for closer inspection. These values were carefully re-checked to ensure the best possible quality data and to retain extreme values where appropriate.

### 3.1.3 Daily Data

**Data Addition.** January 2003 through December 2004 data from NCDC were added to the daily dataset. Extreme values 4" and higher were checked for accuracy by comparing the values with concurrent nearby stations and verifying the data with published sources – NCDC Climatological Data and/or original observation forms as posted on NOAA's "Web Search Store Retrieve Display" (WSSRD) website. There were 61 daily observations  $\geq 4$ " and only one required correction.

**Low Annual Maxima.** HDSC found that some annual maxima were unusually low compared to other annual maxima at nearby stations for a given year. Therefore, as a quality control measure, the two lowest annual maxima were quality controlled at every station in the dataset. Cases where the lowest annual maximum and the second lowest annual maximum were more than 35% different were checked. If crucial months where a high precipitation event occurred at other stations were missing at the given station, then the year at that station was set to missing rather than retain a probably erroneous low annual maximum caused by data sampling. Eleven cases where this occurred were corrected. Typically, the months that were missing were during the hurricane season in these cases. Consideration is being given to modifying the data extraction criteria for this project to screen these types of cases from the annual maximum series, thereby eliminating the need for such a check.

**Hurricane-type Events.** "Hurricane-type events" are defined here as hurricanes, tropical storms, tropical depressions, tropical waves and extra-tropical storms. During

these hurricane-type events, extreme daily observations are sometimes accumulated over several days. In the data extraction process, an accumulated value is distributed evenly over the days of accumulation. Therefore, without information on how the precipitation actually fell over the given days, this may lead to the loss of valuable extreme 1-day precipitation data. For example, station 66-9774 (Villalba 1 E) had a 4 day accumulation of 18.34" during Hurricane Georges ending September 21<sup>st</sup>, 1998. Observations at other stations during Hurricane Georges suggest that most of the precipitation actually fell in one day, which would result in a substantial 1-day annual maximum. To mitigate this issue, software was written to distribute the accumulated values based on spatially distributed ratios from nearby stations.

17 separate hurricane-type events were identified where accumulated values were greater than 5.33", which is the average 1-day mean annual maximum precipitation for all stations. Daily events that exceed the mean are more likely to impact the precipitation frequency analyses. Table 1 lists the events. A total of 139 separate cases of accumulated data were identified at individual stations during these events. Some cases were distributed even though their individual accumulated value was less than 5.33" to be consistent and thorough for a given event.

Since each case may have accumulated values occurring over a different number of days and/or occurring on different days, each case was considered separately in an automated process. Stations that had no accumulated or missing data during the given case were used to appropriately distribute the accumulated data at the target station(s). At each of these stations, ratios of each daily observation to the sum of precipitation on the days of that case (i.e., proportion of total that each day represents) were generated and then spatially interpolated using GRASS's inverse-distance-weighting algorithm. The ratio for each day at the target station was then extracted and applied to generate appropriately distributed daily precipitation values for the accumulated period. This method weights the influence each station has on the distribution at the target station by its relative distance to the target station.

A preliminary examination of 12 randomly-selected cases showed that the procedure was sound and reasonable. The distributed values of each of the 12 cases were manually verified and found to be consistent with nearby stations. As a result we chose not to examine additional cases. Using this objective procedure, the above example, station 66-9774 (Villalba 1 E), resulted in the accumulated 18.34" from Hurricane Georges over the 4 days being distributed as daily values of 0.36", 0.28", 1.59", and 16.11", rather than having an equal distribution of 4.85" on each day. Overall, the means of stations changed by an average of 3.3% ranging from 0.0% to 19.5% as a result of the data corrections.

Table 1. List of hurricane-type storm events in which accumulated observations were distributed where the storm-type is indicated as H (hurricanes), TS (tropical storms), TD (tropical depressions), TW (tropical waves) and EX (extra-tropical storms).

Date	Storm-type	Name
8/12/1956	H	Betsy
8/27/1961	TW	
9/15/1975	TD... TS	Eloise
7/18/1979	TD	
8/30/1979	H	David
9/4/1979	TS	Frederic
12/13/1981	EX	
11/7/1984	TS... H	Klaus
5/18/1985	EX	
10/7/1985	TW	
8/25/1988	TD	
9/18/1989	H	Hugo
7/9/1996	H	Bertha
9/10/1996	H	Hortense
9/22/1998	H	Georges
11/18/1999	H	Lenny
9/15/2004	TS	Jeanne

**Concurrent 1-day/24-hour Check.** HDSC observed during the work to generate the 1-day to 24-hour conversion factors (see Section 3.3, Conversion Factors), that some 24-hour accumulated maximums were significantly higher than their concurrent 1-day counterpart. A 1-day observation is from a gauge measured once each day. At co-located stations it is paired with an hourly gauge that measures once each hour. The 24-hour maximum is obtained by using a moving window to accumulate 1-hour observations from the hourly gauge to 24 hours. It is expected for the 24-hour accumulation to be slightly higher than the 1-day observation because maximum 24-hour amounts seldom fall within a single daily observation period, which is what the conversion factor corrects. However, overly large differences between concurrent monthly 1-day and 24-hour maxima prompted an investigation.

All cases where the 24-hour accumulation was greater than 10" were identified since 10" is a significant maximum likely to have influence in the precipitation frequency analysis. Within this subset, all cases where the percentage difference between the 1-day and 24-hour accumulation were more than 20% were investigated. 20 cases were identified. Half of the cases resulted because the maximum 24-hour accumulation fell over period which bridged between 2 daily observations (e.g., a 12.60" observed as a maximum hourly 24-hour accumulation was in reality measured over 2 days by the daily gauge, 8.88" and 6.26"). However, in the other half, the daily and hourly data were not consistent with the daily data being significantly lower (e.g., a 11.90" 24-hour

accumulation could not be verified by consecutive daily observations, 2.52” and 2.15”). These recorded data were verified in NCDC-published Climatological Data and/or original observation forms as posted on NOAA’s “Web Search Store Retrieve Display” (WSSRD) website.

Nine of the 10 cases in which the daily observations were significantly lower than the hourly accumulations (and 16 of the 20 total cases) occurred during hurricane-type events with high sustained winds. Gauge differences are being investigated to assess whether the high sustained winds led to under-catch, over-catch or other measurement errors in the daily and/or hourly data. These 10 cases were omitted from the calculation of the conversion factor.

Based on observations at nearby stations and storm reports, it is suspected that the hourly data is more reliable in the 10 cases where the daily observations were significantly lower than the hourly accumulations. Therefore, the 1-day observation will be replaced by the 24-hour accumulated observation that has been converted using the inverse of the established conversion factor to a 1-day value for these cases. This will mitigate the discrepancies, since it is unreasonable for co-located daily and hourly stations to be inconsistent.

**Screening for Gaps.** Large gaps (i.e., sequential missing years) in the annual maximum series of stations were screened since it is not possible to guarantee that the two given data segments are from the same population (i.e., same climatology, same rain gauge, same physical environment). Stations with gaps of 5 years or more in the 1-day annual maximum series of stations were flagged and examined on a case-by-case basis. If there were a sufficient number of years (at least 10 years of data) in each data segment, a t-test at 95% confidence level was conducted to assess the statistical integrity of the data record. 34 stations in the Puerto Rico Project area with gaps of 5 years or more were inspected and tested when possible using a t-test. To produce more congruent data records for analysis, station record lengths were truncated where appropriate. The results are as follows:

1. 21 stations with gaps of 5 years or greater passed the t-test and were kept as they are.
2. 7 stations did not have 10 years of data on one side of the gap and so could not be subjected to the t-test. However, based on gap size (generally less than the number of data years before/after the gap), the desire to retain as much data as possible, and the lack of definitive indication that the data were bad, these stations were kept as they are.
3. 2 stations failed the t-test but were kept as they are:
  - a. 66-0158 (Aibonito 1 S) has 70 years of data beginning in 1906 and then a 15 year gap followed by 10 years of data. 10 years is the very minimum required to run the t-test and does not necessarily produce reliable results. Based on the apparent good quality of the data, the capture of Hurricane Georges in 1998 and the desire to retain as much data as possible, it was decided to keep this station as it is.

- b. 66-8822 (San Lorenzo Farm 2 NW) has 28 years of data beginning in 1925 then a 15 year gap followed by 20 years of data. The t-test results are very close to passing. Therefore, in the interest of retaining as much data as possible and without any definitive indication that the data is bad, it was decided to keep this station as it is.
4. 4 stations were truncated because they had very large gaps (greater than 20 years) with 6 years of data or less at the beginning of the record (66-5097, 66-3871, 66-3145, 67-2823)

**Station Corrections.** Several issues related to station location or data were identified and addressed. Daily station 66-0842 (Candelaria Toa Baja) had incorrect coordinates associated with it in the HDSC database. It was corrected to 18.4167 N and 66.2000 W and an elevation of 150 feet. Stations 66-0147 and 66-0152 have same coordinates (17.9667 N and 66.2167 W), same name (Aguirre) and nearly the same elevations (49 and 25 feet, respectively). These two stations are being assessed for the possibility of merging the records together.

Finally, two stations are being investigated for deletion unless local climatological or physical reason can be found for their inconsistent data. Station 66-6017 (Matrullas Dam) may be deleted for the following reasons:

1. 6% of the data are accumulated values occurring in the later part of the record and the data look suspicious.
2. 1932-1954 are hand-entered monthly maxima which allow only for a 1-day duration to be extracted; and since the later part of the record has accumulated data, the 2-day through 4-day durations will not be reliable.
3. There are at least 3 stations in the nearby vicinity to fill in information.
4. The mean, maximum 1-day observation and L-statistics are the lowest of the region suggesting that the data is unreliable.
5. The station data is not consistent with nearby stations as evidenced in the different L-statistics and in a direct comparison of 46 concurrent annual maximums with station 66-9466 (Toro Negro Plant 2), the closest station with concurrent data:
  - a. In 30.4% of the pairs, the annual maxima of 66-6017 were more than 30% lower and 8.7% of the pairs were cases where the annual maximum of 66-9466 was > 10", which suggests an inadequacy in the data of extreme events at 66-6017.
  - b. In only 13% of the pairs, the annual maxima of 66-6017 were actually more than 30% higher than 66-9466 and ALL of these were cases where the annual maximum was only 4" or less.
  - c. Only 23.9% of the cases were within +/- 10% of each other.

Station 66-0849 (Bayney) may be deleted for the following reasons:

1. The data record has 14 years of data, then is missing 25 years of data, and then has only 10 years of data.
2. 1930-1943 are hand-entered monthly maxima leaving only 11 years of complete data. Therefore, this station will not be available for 2-day or longer durations.

3. Station 66-0426 is nearby and is consistent with the overlapping period of 1970-1973 suggesting that there will be no loss of information if station 66-0849 is deleted.

### 3.2 Annual Maximum Series Extraction

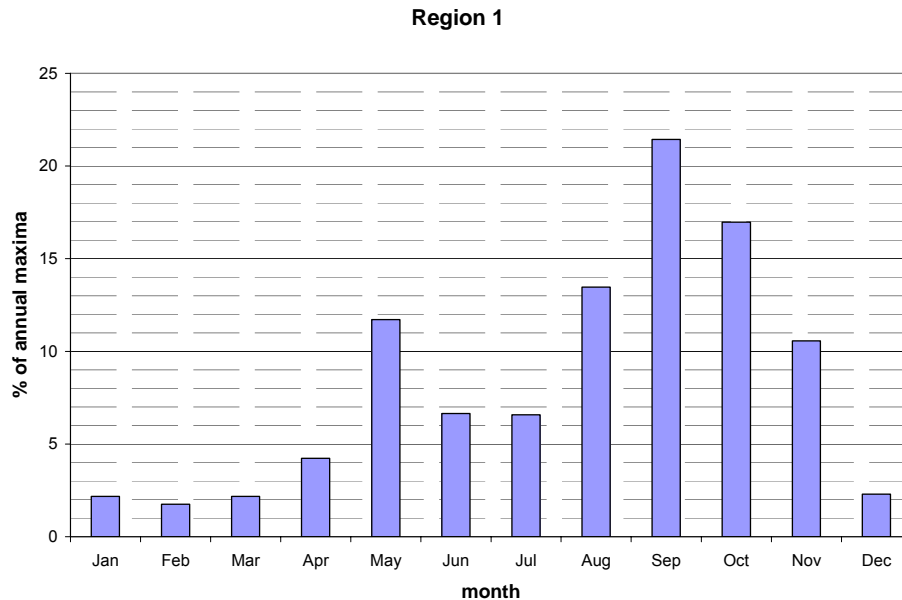
**Minimum Number of Data Years.** Record length may be characterized by the entire period of record or by the number of years of useable data within the total period of record (data years). It was decided to use 20 years as the minimum number of data years for stations to be used in this analysis. Similarly, for the hourly stations a minimum of 15 data years was selected. These minimums allow the inclusion of as many stations as possible and provide a good coverage of the islands, but still maintain the integrity and robustness of the precipitation frequency analysis. These selected minimums are the same as those selected for NOAA Atlas 14 Volume 1 (the semiarid southwestern United States). Because NOAA Atlas 14 Volume 2 (the Ohio River Basin and surrounding states) had significantly more data with longer periods of record, higher minimums were used – 30 data years for daily and 20 data years for hourly.

**Wet Season Criteria.** To accurately extract statistically meaningful values, each year of record must have a sufficient number of data, particularly in an assigned “wet season”. The extraction software requires that at least 50% of the months in the wet season be present to allow an annual maximum to be extracted for a given year. The wet season for each location is defined as the months in which extreme cases were mostly likely to occur. It is assigned by assessing histograms of annual maximum precipitation for each homogeneous region. The histograms show the percentage of annual maxima (derived using no wet season criteria) that occur in each month for all stations in a region. Preliminary wet seasons were assigned to the 6 proposed regions (Figure 1) for the project area based on 1-day data. They are listed in Table 2. Figure 2 shows an example of the histogram for region 1 which has its wet season assigned as May through November (5 through 11). Additional consideration is being given to the particularly extreme nature of precipitation during the hurricane season, September through October. This criteria may be refined for this project.

Table 2. Wet season months for daily regions.

Region	start month	end month
1	5	11
2	4	11
3	4	11
4	4	11
5	1	12
6	4	11

Figure 2. Histogram of annual maximum precipitation for daily region 1.



### 3.3 Conversion Factors

In order to make the daily and hourly data comparable, a conversion is necessary from the daily constrained observation to a 24-hour unconstrained observation (i.e., the accumulated 24-hour observation obtained from 1-hour observations). The conversion factor for this project was computed using ratios of the 2-year quantiles computed from concurrent monthly maxima series (i.e., monthly daily/24-hourly maxima that occurred on the same day) at 18 co-located daily and hourly stations with at least 15 years of concurrent hourly and daily data. A total of 3,217 pairs of monthly maxima were extracted. Monthly maxima time series for concurrent time periods were generated for 24-hour accumulated observations and co-located daily precipitation observations. Ten pairs where the daily observation was inconsistent with hourly accumulation were omitted from the analysis (see Section 3.1.3, Daily Data Collection and Quality Control; Concurrent 1-day/24-hour Check). The series were analyzed separately using L-moments. Ratios of 2-year 24-hour to 2-year 1-day quantiles were then generated and averaged. The conversion factor, 1.208, was the same using different distributions (GNO, GEV, GLO). In the same way, a conversion factor, 1.134, for 2-day to 48-hours was calculated.

These conversion factors were comparable to results using the L-moment results of annual maxima series at these stations. They were slightly lower than factors calculated using a linear regression of daily/24-hourly monthly maxima (or 2-day/48-hourly), 1.220 (or 1.135). The linear regressions were based on the pairs of concurrent data at 18 stations in the project area. However, the linear regression is less reliable because the data are not derived from first order stations and so their concurrent timing is not guaranteed.

The conversion factors for this project are higher than the factors used in NOAA Atlas 14 Volumes 1 and 2 (see Table 3) and in Technical Paper 42 (U.S. Weather Bureau, 1961). However, even though the factors are higher, the ratio of 1-day to 24-hour with 2-day to 48-hour,  $1.208/1.134 = 1.07$ , is comparable with those ratios for NOAA Atlas 14 Volume 1 (Semiarid Southwest), 1.11, and NOAA Atlas 14 Volume 2 (Ohio River Basin), 1.09.

Table 3. Conversion factors for constrained to unconstrained observations.

Project	Conversion Factors		Ratio of Factors (1d24hr/2d48hr)
	1-day to 24-hour	2-day to 48-hour	
NOAA Atlas 14 Vol. 3 (Puerto Rico and the U.S. Virgin Islands)	1.208	1.134	1.07
NOAA Atlas 14 Vol. 1 (Semiarid Southwestern United States)	1.143	1.03	1.11
NOAA Atlas 14 Vol. 2 (Ohio River Basin and Surrounding States)	1.134	1.04	1.09
Technical Paper 42	1.13	N/A	

### 3.4 L-moments and Regionalization

Preliminary L-moments were run on the 24-hour and 1-hour data. L-moment statistics (coefficient of L-variation, L-skewness, and L-kurtosis) can be used to parameterize the probability distributions at each station. Coefficient of L-variation provides a measure of dispersion. L-skewness is a measure of symmetry. L-kurtosis is a measure of peakedness. Figures 3 through 6 show the coefficient-of-L-variation (L-CV) and the L-skewness (L-CS) for the 24-hour and 1-hour station data in the project area. It should be kept in mind that the both the 24-hour and 1-hour datasets continue to undergo quality control so these results may change.

It was decided to pursue new regions for the following reasons:

1. The original 7 regions were not yet substantiated by extreme precipitation patterns and current investigations suggest that they may be inadequate to depict extreme precipitation patterns.
2. There is a disconnect between those original regions and mean annual maxima, maximum observed, and the L-statistic pattern that suggested additional research was advisable.
3. Original region 3 was discordant and seemed particularly difficult to refine given observed spatial patterns.
4. Other studies of regionalization of Puerto Rico which objectively based regions on climate variables (Carter and Elsner, 1996; Carter et. al, 1997 ; Malmgren and Winter, 1999; Colon, 1966) were similar in pattern to mean annual maxima and maximum observed values (as well as L-statistics).

Figure 1 at the beginning of this report depicts the new regions as groupings of stations. The new regions are based on mean annual maximum precipitation, maximum observed precipitation, mean annual precipitation, climate divisions from the San Juan NWS, HDSC's original regions, station data characteristics and previous studies. The designated regions are confirmed by statistical homogeneity tests and other checks. In particular, the heterogeneity measure,  $H_1$ , tests between-station variations in sample L-moments for a group of stations with what would be expected for a homogeneous region based on coefficient of L-variation (Hosking and Wallis, 1997). An  $H_1$  measure greater than 2 ( $H_1 > 2$ ) indicates heterogeneity and  $H_1 < 2$  indicates homogeneity. As suggested in Hosking and Wallis (1997), adjustments of regions, such as moving stations from one region to another or subdividing a region, are being made to reduce heterogeneity. Table 4 shows the preliminary  $H_1$  results for these regions. Currently, only region 2 is heterogeneous. Adjustments to the regions will also be considered based on local climate knowledge and expertise.

Figure 3. Coefficient of L-variation station statistics for 24-hour data.

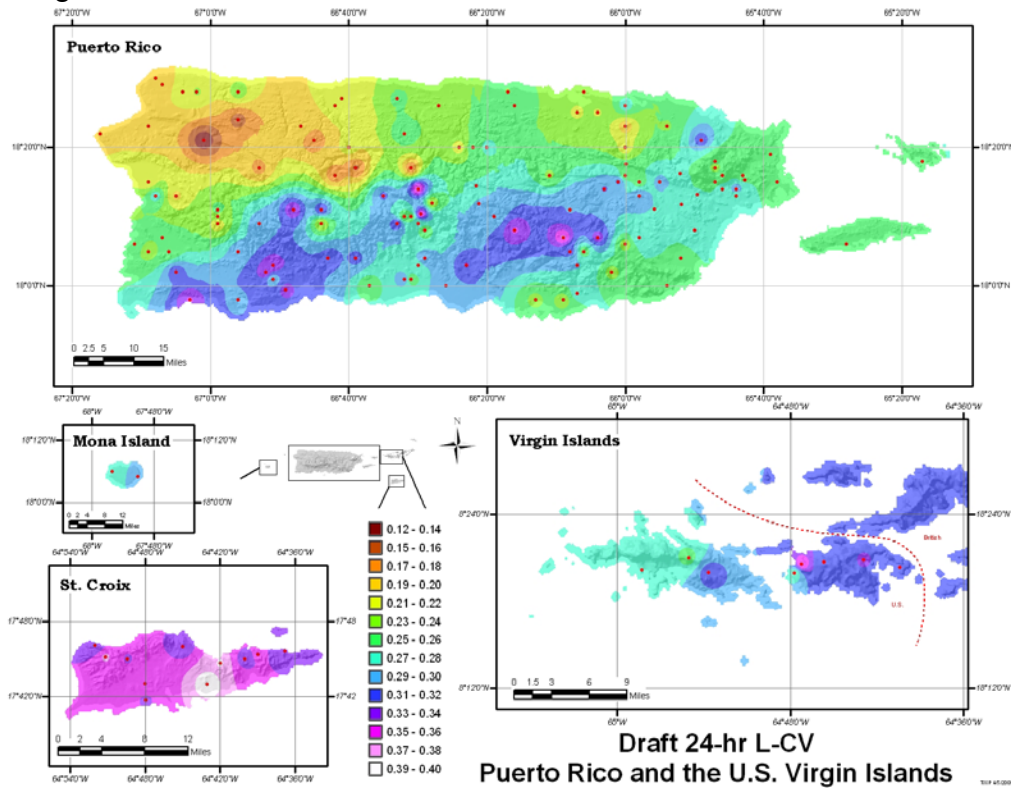


Figure 4. Coefficient of L-skewness station statistics for 24-hour data.

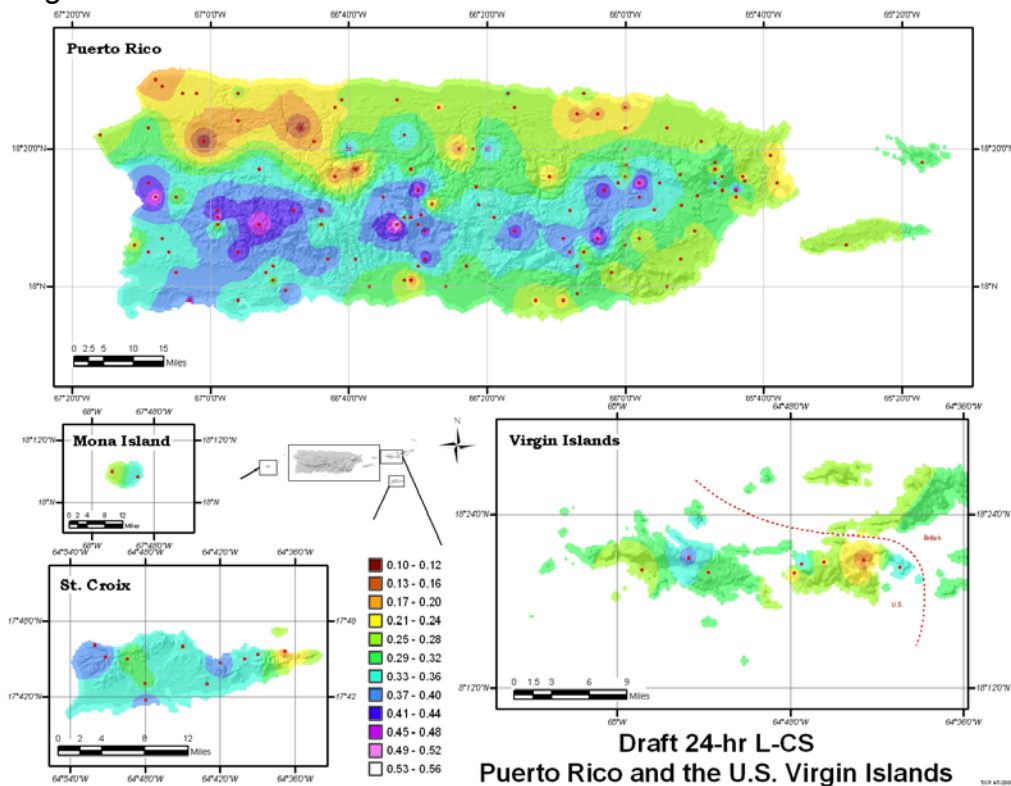


Figure 5. Coefficient of L-variation station statistics for 1-hour data.

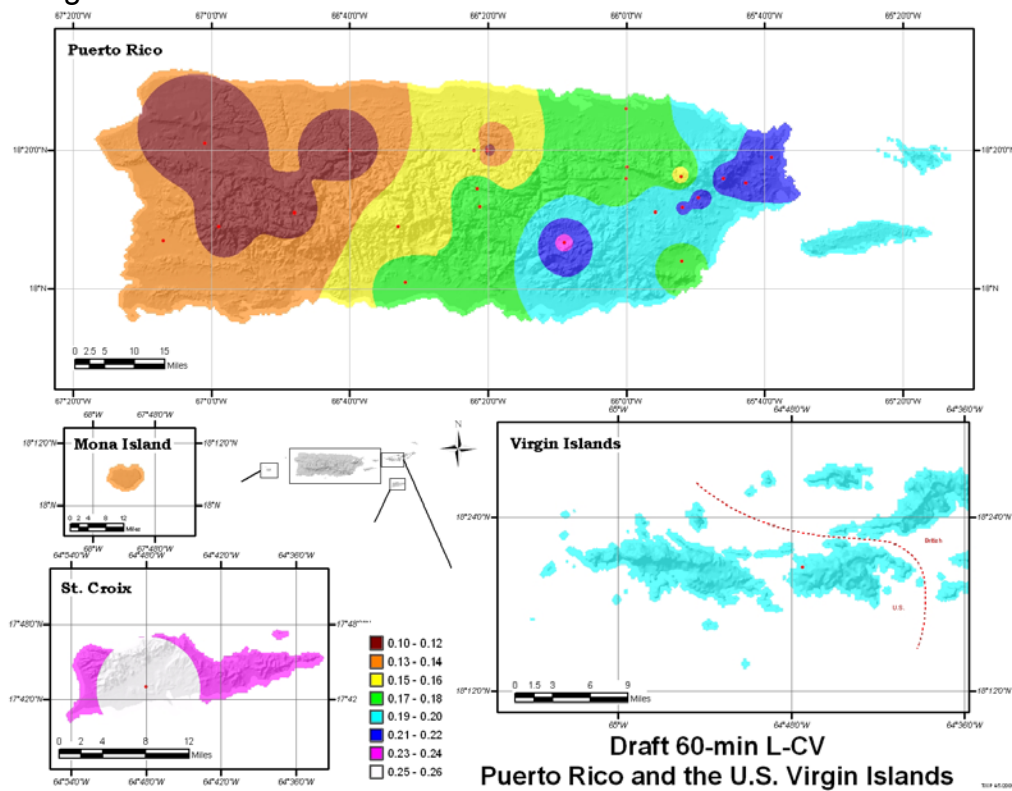


Figure 6. Coefficient of L-skewness station statistics for 1-hour data.

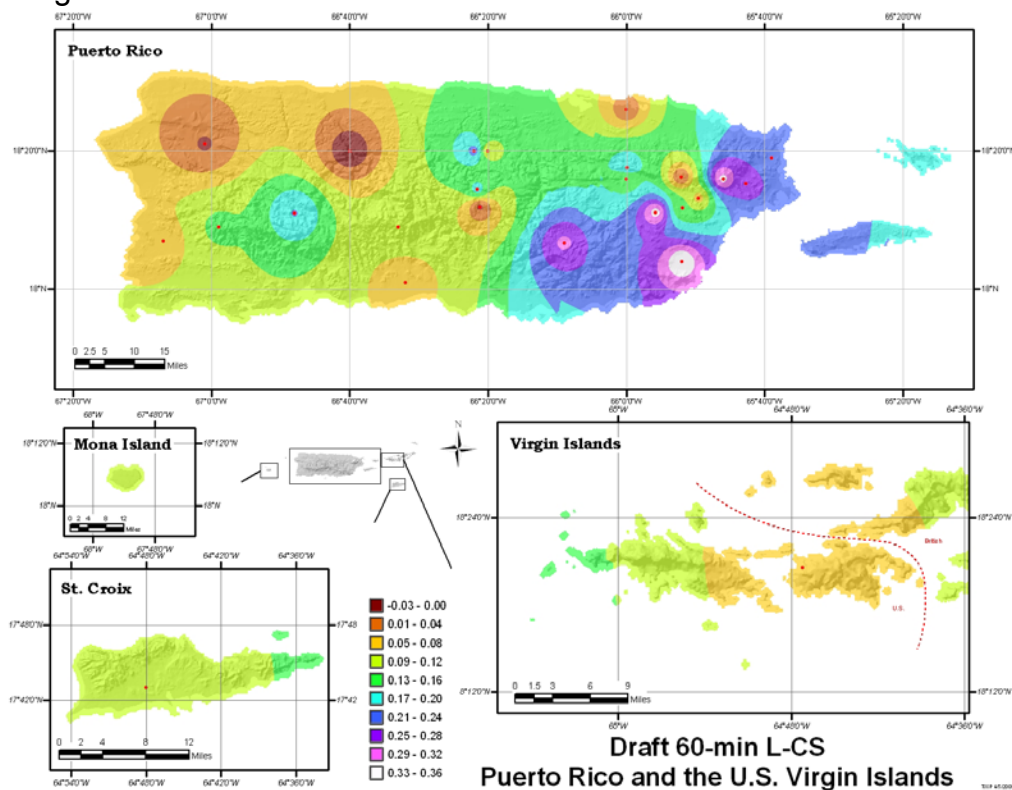


Table 4. Preliminary homogeneity/heterogeneity results for initial regions in Figure 1.

region	# daily stations	# hourly stations	# total stations	H1
1	27	4	31	1.40
<b>2</b>	<b>18</b>	<b>2</b>	<b>20</b>	<b>2.97</b>
3	20	3	23	1.11
4	25	9	34	1.29
5	17	5	22	-0.55
6	23	2	25	0.19
<b>total</b>	<b>130</b>	<b>25</b>	<b>155</b>	

### 3.5 Temporal Distributions

Temporal distributions of heavy precipitation will be provided for use with precipitation frequency for 6, 12 , 24- and 96-hour durations. The temporal distributions are expressed in probabilistic terms as cumulative percentages of precipitation and duration at various percentiles. The method of obtaining temporal distributions for this project is the same as described for NOAA Atlas 14 Volume 1 in Appendix A.1 (Bonnin et al., 2004).

Preliminary temporal distributions for the 24-hour duration for the project area as a whole were calculated and are presented in Figure 7. The data were also subdivided into quartiles based on where in the distribution the most precipitation occurred in order to provide more specific information on the varying distributions that were observed. Figure 8 depicts preliminary temporal distributions for each quartile for the 24-hour duration. Table 5 lists the number and proportion of cases in each quartile for the Island as a whole. The preliminary results are reasonable and seem consistent with previous results for NOAA Atlas 14 Volumes 1 and 2. Figure 9 shows an over-laid comparison of the results for this project compared to Volume 2 (Ohio River basin and surrounding states).

Figure 7.

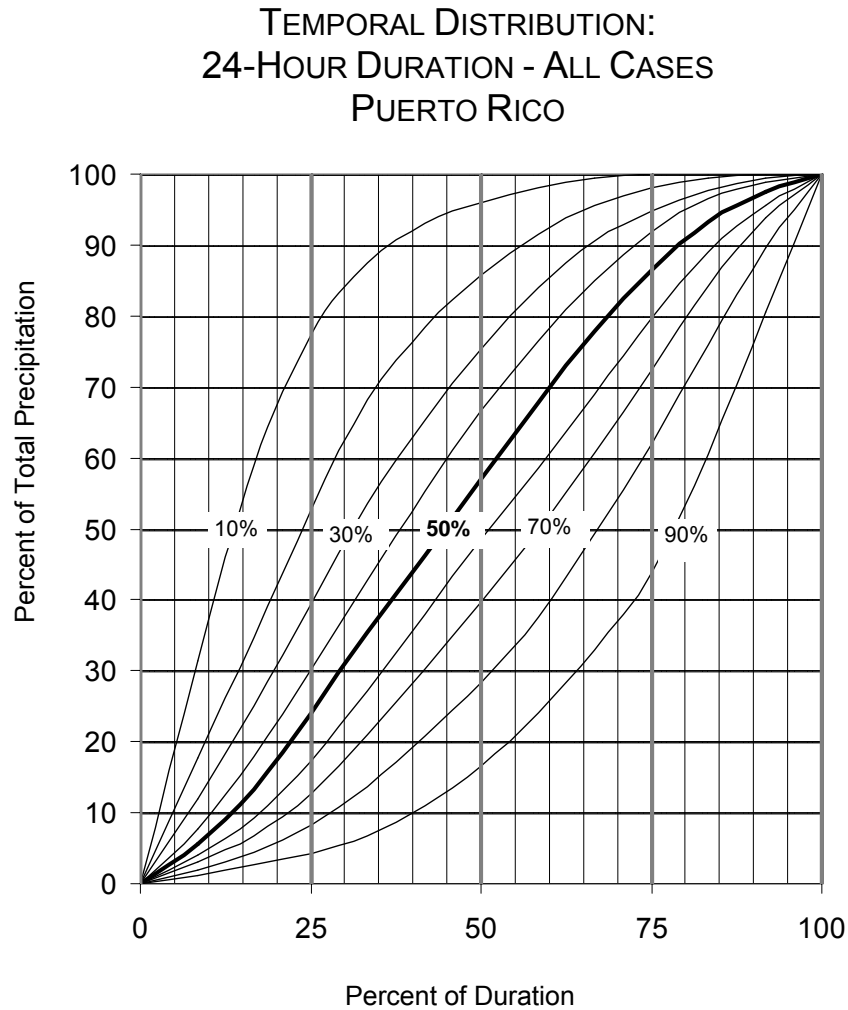


Table 5. Numbers and proportion of cases in each quartile and temporal distribution associated with NOAA Atlas 14 Volume 3.

	1 <sup>st</sup> Quartile	2 <sup>nd</sup> Quartile	3 <sup>rd</sup> Quartile	4 <sup>th</sup> Quartile	Total number of cases
24-hour	135 (32%)	112 (26%)	95 (22%)	84 (20%)	426

Figure 8.

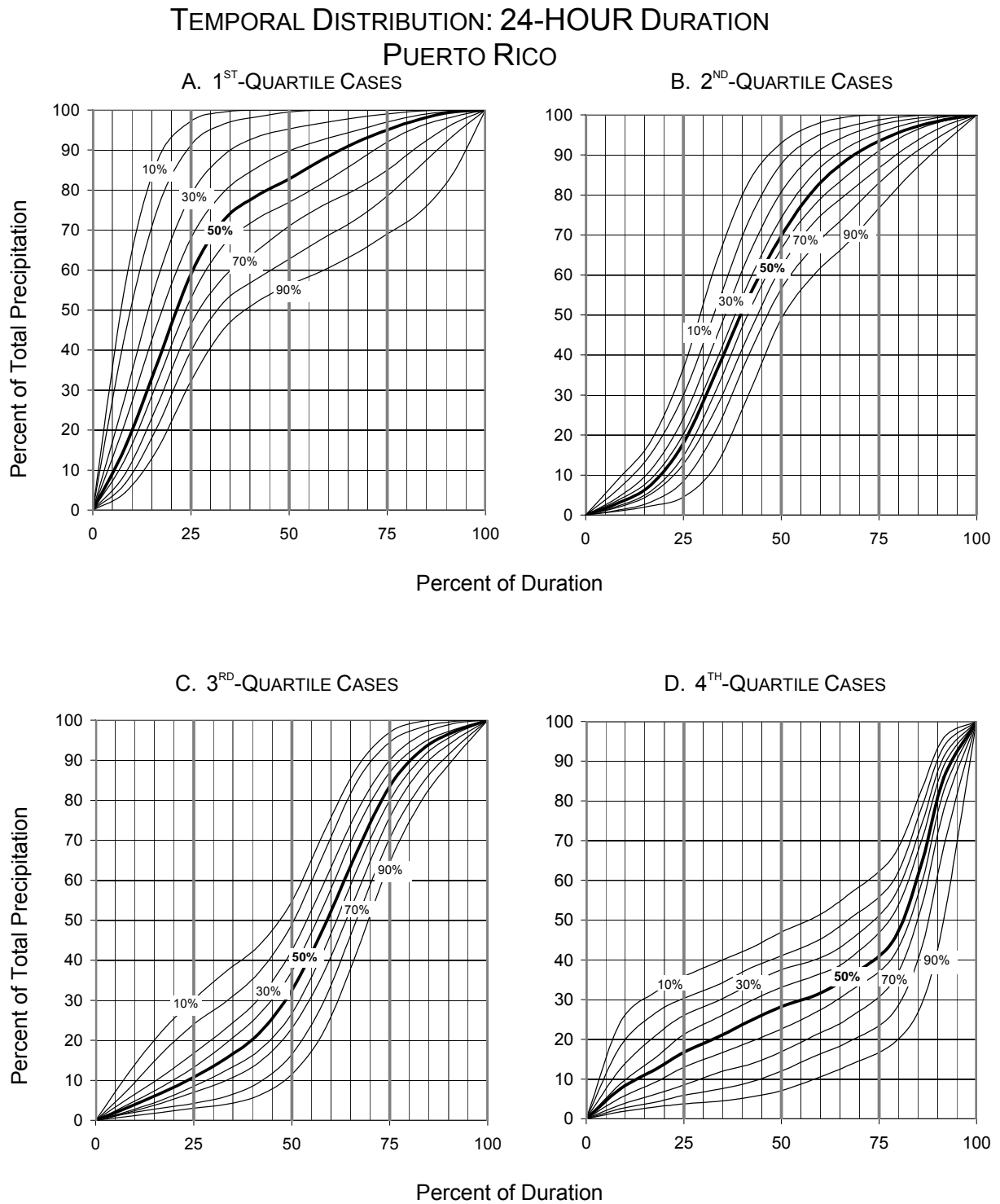
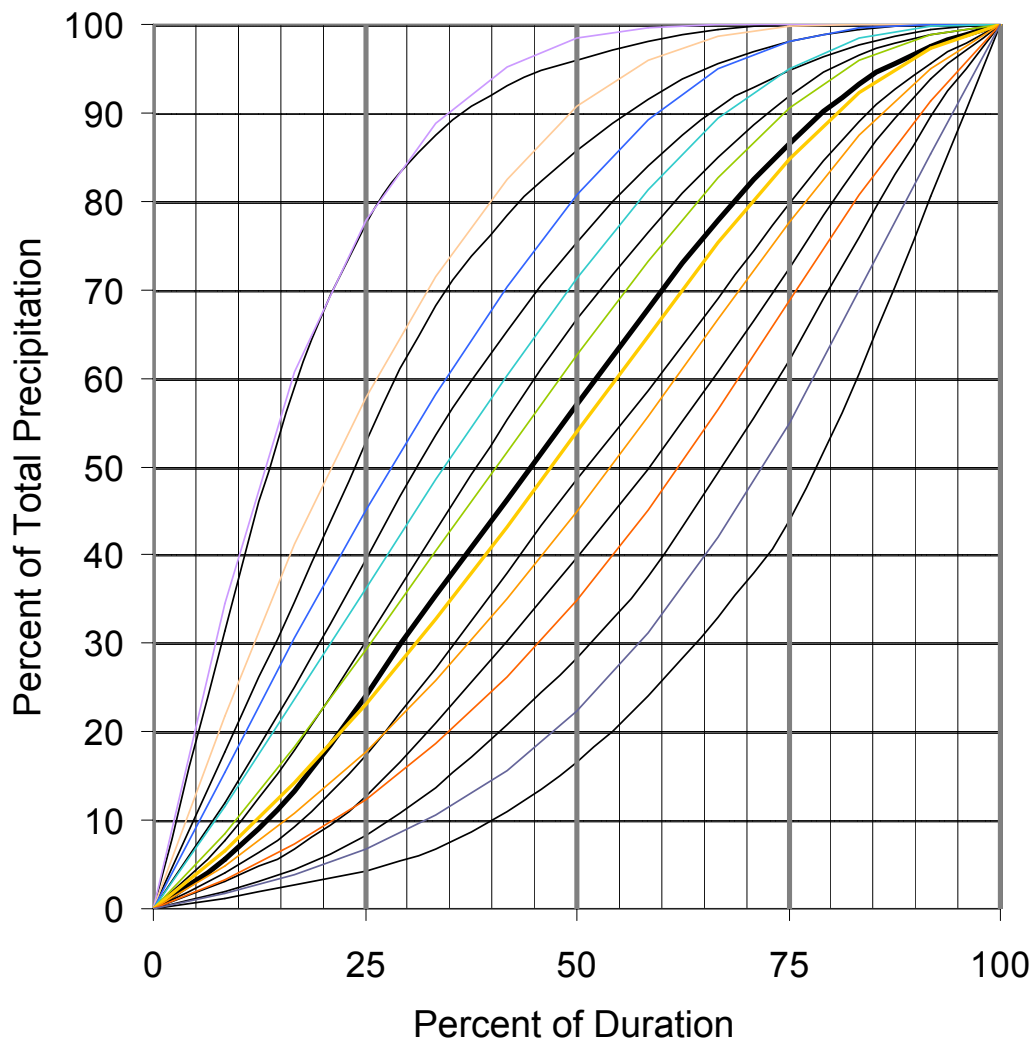


Figure 9.

TEMPORAL DISTRIBUTION:  
24-HOUR DURATION - ALL CASES  
PUERTO RICO (BLACK) VS. OHIO RIVER BASIN (COLOR)



### 3.6 Spatial Interpolation

Select maps from Technical Paper 42 (U.S. Weather Bureau, 1961) were digitized to allow objective comparisons with this project's results. The isohyets from the 100-year 6-hour and 100-year 24-hour maps were digitized into a geographic information system (GIS), and then converted into high-resolution grids. Although isohyets are available for Puerto Rico and St. Croix, the other islands (St. Thomas, Mona, Culebra and Vieques) only had point estimates in Technical Paper 42. The point estimates on these islands were digitized and spatially distributed to a grid using an inverse-distance-weighting algorithm. The grids from each of the islands were then patched together to make one complete grid for the entire Puerto Rico-U.S. Virgin Island region.

### 3.7 PFDS

The Precipitation Frequency Data Server (PFDS), the on-line portal for all NOAA Atlas 14 deliverables and information, under went several important changes. The most noticeable change was the re-organization of the GIS Download page. To ease confusion and increase usability, this page, which was originally multiple pages long, was split into four separate web pages titled: Time Series, Maps, GIS Data and Version Numbers. As a result, the left sidebar items changed accordingly. Furthermore, buttons to these pages were added to the header of the state specific pages to allow easier navigation.

Other changes include:

1. Continued to update the PFDS Performance and Stats page on a monthly basis (see below).
2. Added a link to Customer Survey results so users could view the categorical feedback we're getting.
3. A few frequently asked questions (FAQ) were added to the FAQ page.

HDSC continuously monitors the hits, integrity and performance of the PFDS, which continues to receive an increasing number of hits per month. The graph (Figure 10) below summarizes the number of individual data inquiries made since January 2004, while the map (Figure 11) indicates the locations of inquiries during the past quarter.

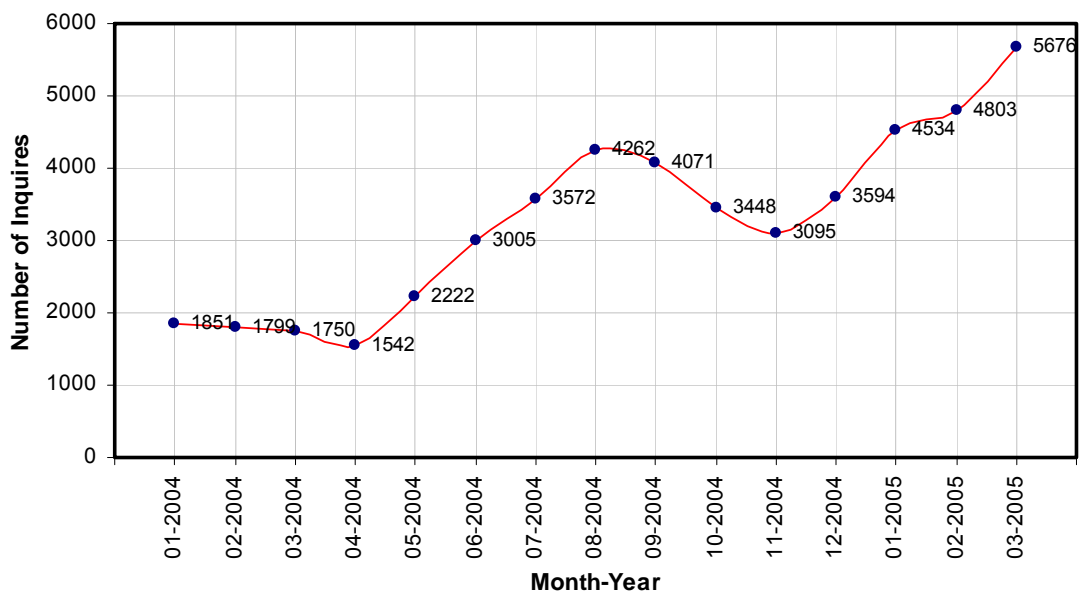


Figure 10: Number of individual PFDS data inquiries per month.

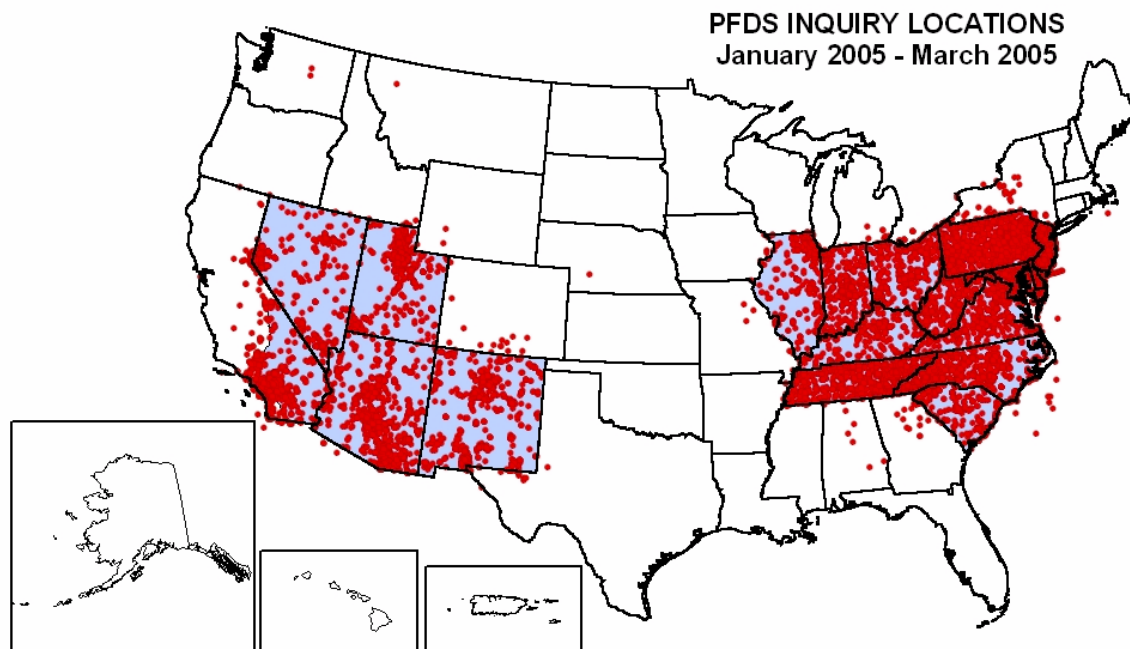


Figure 11: Map of 15,013 PFDS data inquiry locations during the period January-March 2005.

### **3.8 Areal Reduction Factors**

Work continues in the development of geographically-fixed Areal Reduction Factor (ARF) curves for basin area sizes of 10 to 400 square miles. Progress has been slow due to difficulties in completing the analysis software. Development and testing of software from the procedure described in NOAA Technical Report NWS 24 is 90% completed. Modifications have been made in the ARF sites with respect to which stations are being used. This is being done to ensure distances between stations used in the calculations are appropriate. In particular, if a station is more than 80 miles away from any other station for that study site, then that station will not be used in that site's ARF analysis.

Two statistical and objective testing procedures, the sign test (Himmelblau, 1970) and a modified "student t" test (Siegel, 1961) will be used in testing the differences in the ARF curves generated from the various sites. A third objective statistical test for testing differences in the ARF curves is currently being investigated. These procedures were also used in an ARF study performed by Bell (1976).

## **4. Issues**

### **4.1 Recent and Upcoming Presentations**

On March 30 and 31, Geoff Bonnin traveled to Puerto Rico to discuss progress on updating precipitation frequency estimates for Puerto Rico and the U.S. Virgin Islands. In a series of meetings, he met with the NOAA's San Juan Weather Forecast Office (WFO), Puerto Rico and the Commonwealth of Puerto Rico's Department of Natural Resources and Minerals in San Juan. Other interested parties attended such as U.S. Geological Survey, U.S. Army Corps of Engineers, U.S. Environmental Protection Agency and the U.S. Department of Agriculture. This facilitated an important exchange regarding the local climate and feedback on the regions to be used in this project.

HDSC has submitted a poster paper entitled "NOAA Atlas 14, the new precipitation frequency atlas for the United States" for the European Geosciences Union General Assembly meeting on April 24 – 29, 2005 in Vienna, Austria. Due to funding and schedule considerations, the poster will be presented by others from the Office of Hydrologic Development already attending the meeting.

### **4.2 1-year Precipitation Frequency**

HDSC has been approached by the State of Maryland State Highway Administration (MDSHA) to calculate and include the 1-year average recurrence interval (ARI) precipitation frequency estimates for NOAA Atlas 14 Volume 2. Discussions are being held with MDSHA on funding and contractual mechanisms and the areas to be covered.

## **5. Projected Schedule and Remaining Tasks**

The following list provides a tentative schedule with completion dates. Brief descriptions of tasks to be worked on are also included in this section.

- Data Collection and Quality Control [April 2005]
- Trend Analysis [April 2005]
- Temporal Distributions of Extreme Rainfall [April 2005]
- L-Moment Analysis/Frequency Distribution [April 2005]
- Peer Review of Spatially Interpolated Point Estimates [July 2005]
- Spatial Interpolation of Grids [August 2005]
- Precipitation Frequency Maps [September 2005]
- Web Publication [August 2005]
- Spatial Relations (Areal Reduction Factors) [May 2005]

### **5.1 Data Collection and Quality Control**

During the next quarter, quality control of the updated n-minute dataset will be completed. Conversion factors and n-minute ratios will be calculated.

### **5.2 L-Moment Analysis/Frequency Distribution**

A comprehensive L-moment statistical analysis and regionalization will be completed on the 24-hour and 60-minute durations.

### **5.3 Trend Analysis and Temporal Distributions**

Once the data have been quality controlled, an analysis for trends in the annual maximum time series and an analysis of the hourly temporal distributions of heavy rainfall can be completed.

### **5.4 Areal Reduction Factors (ARF)**

Computations for the ARF curves will be completed in the next quarter for 14 areas. The resulting curves will be tested for differences to determine if a single set of ARF curves is applicable to the entire U.S. or whether curves vary by region.

## References

- Arkell, R.E., and F. Richards, 1986: Short duration rainfall relations for the western United States, *Conference on Climate and Water Management-A Critical Era and Conference on the Human Consequences of 1985's Climate*, August 4-7, 1986. Asheville, NC.
- Bell, F.C., 1976: The areal reduction factors in rainfall frequency estimation, *Natural Environmental Research Council (NERC), Report No. 35*, Institute of Hydrology, Wallingford, U.K. 25pp.
- Bonnin, G., D. Todd, T. Parzybok, B. Lin, D. Riley, and M. Yekta, 2004: Precipitation frequency atlas of the United States. *NOAA Atlas 14 Volume 1*, Silver Spring, Maryland. <http://hdsc.nws.noaa.gov/hdsc/>.
- Carter, M. M. and J. B. Elsner, 1996: Convective Rainfall Regions of Puerto Rico, *International Journal of Climatology*, **16**, pp. 1033-1043.
- Carter, M. M., S. Bennett, and J.B. Elsner, 1997: Monthly Rainfall Climatology for Puerto Rico, *NWS Southern Topics February 1997*.
- Colon, J, 1966: On the Mechanisms for Production of Rainfall in Puerto Rico, *Weather Bureau Technical Memorandum 15*, San Juan, Puerto Rico.
- Frederick, R.H. and J.F. Miller, 1979: Short Duration Rainfall Frequency Relations for California, Third Conference on Hydrometeorology, August 20-24, 1979. Bogata Columbia.
- Frederick, R.H., V.A. Myers and E.P. Auciello, 1977: Five- to 60-minute precipitation frequency for the eastern and central United States, *NOAA Technical Memo. NWS HYDRO-35*, Silver Spring, MD, 36 pp.
- Hershfield, D.M., 1961: Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years, *Weather Bureau Technical Paper No. 40*, U.S. Weather Bureau. Washington, D.C., 115 pp.
- Himmelblau, D.M., 1970: "Sign Test for Median Difference in Paired Observations", *Process Analysis by Statistical Methods*, page 68.
- Hosking, J.R.M. and J.R. Wallis, 1997: *Regional frequency analysis, an approach based on L-moments*, Cambridge University Press, 224 pp.
- Huff, F. A., 1990: Time Distributions of Heavy Rainstorms in Illinois, *Illinois State Water Survey*, Champaign, 173, 17pp.

Institution of Engineers, Australia, 1987: *Australian Rainfall and Runoff, 3rd Edition*, the Institution of Engineers, Australia. Canberra.

Lin, B. and L.T. Julian, 2001: Trend and shift statistics on annual maximum precipitation in the Ohio River Basin over the last century. Symposium on Precipitation Extremes: Prediction, Impacts, and Responses, 81st AMS annual meeting. Albuquerque, New Mexico.

Malgren, B., and A. Winter, 1999: Climate Zonation in PR based on Principal Component Analysis and an Artificial Neural Network, *Journal of Climate*, **12**, pp. 977-985.

Miller, J.F., 1964: Two- to ten-day precipitation for return periods of 2 to 100 years in the contiguous United States, *Technical Paper No. 49*, U.S. Weather Bureau and U.S. Department of Agriculture, 29 pp.

Miller, J.F., 1965: Two- to Ten-Day Rainfall for Return Periods of 2 to 100 Years in Puerto Rico and the Virgin Islands, *Technical Paper No. 53*, U.S. Weather Bureau and U.S. Department of Agriculture, 35 pp.

Miller, J.F., R.H. Frederick and R.J. Tracy, 1973: Precipitation-frequency atlas of the western United States, *NOAA Atlas 2*, 11 vols., National Weather Service, Silver Spring, MD.

Myers, V.A. and R.M. Zehr, 1980: A Methodology for Point-to-Area Rainfall Frequency Ratios, *NOAA Technical Report NWS 24*, Office of Hydrology, National Weather Service, Silver Spring, MD.

Spiegel, M.R., 1961: "Tests of Significance Involving Sample Differences", *Theory and Problems of Statistics*, pages 170-171.

U.S. Weather Bureau, 1961: Generalized Estimates of Probable Maximum Precipitation and Rainfall-Frequency Data for Puerto Rico and Virgin Islands for Areas to 400 Square Miles, Durations to 24 Hours and Return Periods from 1 to 100 Years, *Technical Paper 42*, Washington, DC, 94pp.